

# 17



1600

## RAW SEQUENCE LISTING

DATE: 04/04/2002

PATENT APPLICATION: US/09/389,537B

TIME: 10:12:21

Input Set : A:\DIVER1240-3.ST25.txt

Output Set: N:\CRF3\04032002\I389537B.raw

ENTERED

P.S

2 <110> APPLICANT: DIVERSA CORPORATION  
 3 WARREN, Patrick  
 4 SWANSON, Ronald  
 6 <120> TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES  
 8 <130> FILE REFERENCE: DIVER1240-3  
 10 <140> CURRENT APPLICATION NUMBER: US 09/389,537B  
 11 <141> CURRENT FILING DATE: 1999-09-02  
 13 <150> PRIOR APPLICATION NUMBER: US 08/646,590  
 14 <151> PRIOR FILING DATE: 1996-05-08  
 16 <150> PRIOR APPLICATION NUMBER: US 08/599,171  
 17 <151> PRIOR FILING DATE: 1996-02-09  
 19 <160> NUMBER OF SEQ ID NOS: 40  
 21 <170> SOFTWARE: PatentIn version 3.0  
 23 <210> SEQ ID NO: 1  
 24 <211> LENGTH: 52  
 25 <212> TYPE: DNA  
 26 <213> ORGANISM: Artificial sequence  
 28 <220> FEATURE:  
 29 <223> OTHER INFORMATION: Primer for PCR  
 31 <400> SEQUENCE: 1  
 32 ccgagaattc attaaagagg agaaattaac tatgattgaa gaccctatgg ac 52  
 35 <210> SEQ ID NO: 2  
 36 <211> LENGTH: 29  
 37 <212> TYPE: DNA  
 38 <213> ORGANISM: Artificial sequence  
 40 <220> FEATURE:  
 41 <223> OTHER INFORMATION: Primer for PCR  
 43 <400> SEQUENCE: 2  
 44 cttggactct cttcacgatt tctagaagc 29  
 47 <210> SEQ ID NO: 3  
 48 <211> LENGTH: 52  
 49 <212> TYPE: DNA  
 50 <213> ORGANISM: Artificial sequence  
 52 <220> FEATURE:  
 53 <223> OTHER INFORMATION: Primer for PCR  
 55 <400> SEQUENCE: 3  
 56 ccgagaattc attaaagagg agaaattaac tatggacagg cttgaaaaag ta 52  
 59 <210> SEQ ID NO: 4  
 60 <211> LENGTH: 31  
 61 <212> TYPE: DNA  
 62 <213> ORGANISM: Artificial sequence  
 64 <220> FEATURE:  
 65 <223> OTHER INFORMATION: Primer for PCR

## RAW SEQUENCE LISTING

DATE: 04/04/2002

PATENT APPLICATION: US/09/389,537B

TIME: 10:12:21

Input Set : A:\DIVER1240-3.ST25.txt

Output Set: N:\CRF3\04032002\I389537B.raw

```

67 <400> SEQUENCE: 4
68 aagaatctct tcgaatcgac ttctagaagg c 31
71 <210> SEQ ID NO: 5
72 <211> LENGTH: 52
73 <212> TYPE: DNA
74 <213> ORGANISM: Artificial sequence
76 <220> FEATURE:
77 <223> OTHER INFORMATION: Primer for PCR
79 <400> SEQUENCE: 5
80 ccgacaattg attaaagagg agaaattaac tatgtgggaa ttagacccta aa 52
83 <210> SEQ ID NO: 6
84 <211> LENGTH: 30
85 <212> TYPE: DNA
86 <213> ORGANISM: Artificial sequence
88 <220> FEATURE:
89 <223> OTHER INFORMATION: Primer for PCR
91 <400> SEQUENCE: 6
92 tcgaactttt tctccacatc cctaggaggc 30
95 <210> SEQ ID NO: 7
96 <211> LENGTH: 52
97 <212> TYPE: DNA
98 <213> ORGANISM: Artificial sequence
100 <220> FEATURE:
101 <223> OTHER INFORMATION: Primer for PCR
103 <400> SEQUENCE: 7
104 ccgacaattg attaaagagg agaaattaac tatgacatac ttaatgaaca at 52
107 <210> SEQ ID NO: 8
108 <211> LENGTH: 31
109 <212> TYPE: DNA
110 <213> ORGANISM: Artificial sequence
112 <220> FEATURE:
113 <223> OTHER INFORMATION: Primer for PCR
115 <400> SEQUENCE: 8
116 gaactttccc tgaagagtat ttctagaagg c 31
119 <210> SEQ ID NO: 9
120 <211> LENGTH: 52
121 <212> TYPE: DNA
122 <213> ORGANISM: Artificial sequence
124 <220> FEATURE:
125 <223> OTHER INFORMATION: Primer for PCR
127 <400> SEQUENCE: 9
128 ccgagaattc attaaagagg agaaattaac tatgcggaaa ctggccgagc gg 52
131 <210> SEQ ID NO: 10
132 <211> LENGTH: 31
133 <212> TYPE: DNA
134 <213> ORGANISM: Artificial sequence
136 <220> FEATURE:
137 <223> OTHER INFORMATION: Primer for PCR
139 <400> SEQUENCE: 10

```

## RAW SEQUENCE LISTING

DATE: 04/04/2002

PATENT APPLICATION: US/09/389,537B

TIME: 10:12:21

Input Set : A:\DIVER1240-3.ST25.txt

Output Set: N:\CRF3\04032002\I389537B.raw

```

140 aactagcttc gccgtgaaat tcctaggagg c 31
143 <210> SEQ ID NO: 11
144 <211> LENGTH: 52
145 <212> TYPE: DNA
146 <213> ORGANISM: Artificial sequence
148 <220> FEATURE:
149 <223> OTHER INFORMATION: Primer for PCR
151 <400> SEQUENCE: 11
152 ccgacaattg attaaagagg agaaattaac tatgtgcggg atagtcggat ac 52
155 <210> SEQ ID NO: 12
156 <211> LENGTH: 31
157 <212> TYPE: DNA
158 <213> ORGANISM: Artificial sequence
160 <220> FEATURE:
161 <223> OTHER INFORMATION: Primer for PCR
163 <400> SEQUENCE: 12
164 ttttgccagt gccaccttat ttctagaagg c 31
167 <210> SEQ ID NO: 13
168 <211> LENGTH: 52
169 <212> TYPE: DNA
170 <213> ORGANISM: Artificial sequence
172 <220> FEATURE:
173 <223> OTHER INFORMATION: Primer for PCR
175 <400> SEQUENCE: 13
176 ccgacaattg attaaagagg agaaattaac tatgataccc cagaggatta ag 52
179 <210> SEQ ID NO: 14
180 <211> LENGTH: 31
181 <212> TYPE: DNA
182 <213> ORGANISM: Artificial sequence
184 <220> FEATURE:
185 <223> OTHER INFORMATION: Primer for PCR
187 <400> SEQUENCE: 14
188 agggaaagtt cgagagaaat ttctagaagg c 31
191 <210> SEQ ID NO: 15
192 <211> LENGTH: 52
193 <212> TYPE: DNA
194 <213> ORGANISM: Artificial sequence
196 <220> FEATURE:
197 <223> OTHER INFORMATION: Primer for PCR
199 <400> SEQUENCE: 15
200 ccgagaattc attaaagagg agaaattaac tatgaagccg tacgctaaat at 52
203 <210> SEQ ID NO: 16
204 <211> LENGTH: 31
205 <212> TYPE: DNA
206 <213> ORGANISM: Artificial sequence
208 <220> FEATURE:
209 <223> OTHER INFORMATION: Primer for PCR
211 <400> SEQUENCE: 16
212 acctagtggag gacacataat ctctagaagg c 31

```

## RAW SEQUENCE LISTING

DATE: 04/04/2002

PATENT APPLICATION: US/09/389,537B

TIME: 10:12:21

Input Set : A:\DIVER1240-3.ST25.txt

Output Set: N:\CRF3\04032002\I389537B.raw

215 &lt;210&gt; SEQ ID NO: 17

216 &lt;211&gt; LENGTH: 1245

217 &lt;212&gt; TYPE: DNA

218 &lt;213&gt; ORGANISM: Aquifex

220 &lt;400&gt; SEQUENCE: 17

221	atgattgaag	accctatgga	ctgggctttt	ccgaggataa	agagactgcc	tcagtatgtc	60
223	ttctctctcg	ttaacgaact	caagtacaag	ctaaggcgtg	aaggcgaaga	tgtagtggtat	120
225	cttggtatgg	gcaatcctaa	catgcctcca	gcaaagcaca	taatagataa	actctgcgaa	180
227	gtggctcaaa	agccgaacgt	tcacggatat	tctgcgtcaa	ggggcatacc	aagactgaga	240
229	aaggctatat	gtaacttcta	cgaagaaagg	tacggagtga	aactcgaccc	tgagagggag	300
231	gctatactaa	caatcgggtg	aaaggaaggg	tattctcatt	tgatgcttgc	gatgatattct	360
233	ccgggtgata	cggtaatagt	tcctaatacc	acctatccta	ttcactatta	cgctcccata	420
235	attgcaggag	gggaagttca	ctcaataccc	cttaacttct	cggacgatca	agatcatcag	480
237	gaagagtttt	taaggaggct	ttacgagata	gtaaaaaccg	cgatgccaaa	acccaaggct	540
239	gtcgtcataa	gctttcctca	caatccaacg	accataacgg	tagaaaagga	cttttttaaa	600
241	gaaatagtta	agtttgcaaa	ggaacacggt	ctctggataa	tacacgattt	tgcgtatgcg	660
243	gatatagcct	ttgacgggta	caagccccc	tcaatactcg	aaatagaagg	tgctaaagac	720
245	gttgcggttg	agctctactc	catgtcaaa	ggcttttcaa	tgccgggctg	gagggtagcc	780
247	tttgcgttg	gaaacgaaat	actcataaaa	aaccttgac	acctcaaaag	ctacttggtat	840
249	tacggtatat	ttactcccat	acaggtggcc	tctattatcg	cattagagag	cccctacgaa	900
251	atcgtggaaa	aaaccgcaaa	ggtttaccaa	aaaagaagag	acgttctggt	ggaaggggta	960
253	aacaggctcg	gctggaaaag	aaaaaaacct	aaggctacca	tgctcgtctg	ggcaaagatt	1020
255	cccgaatgga	taaatatgaa	ctctctggac	ttttccttgt	tcctcctaaa	agaggcgaa	1080
257	gttgcggtat	ccccgggtgt	gggctttggt	cagtacggag	aggggtacgt	aaggtttgca	1140
259	cttgtagaaa	atgaacacag	gatcagacag	gctataaggg	gaataaggaa	agccttcaga	1200
261	aaactccaga	aggagaggaa	acttgaacct	gagagaagtg	cttaa		1245

264 &lt;210&gt; SEQ ID NO: 18

265 &lt;211&gt; LENGTH: 1122

266 &lt;212&gt; TYPE: DNA

267 &lt;213&gt; ORGANISM: Aquifex

269 &lt;400&gt; SEQUENCE: 18

270	atggacaggc	ttgaaaaagt	atcacccttc	atagtaatgg	atatacctagc	tcaggcccag	60
272	aagtacgaag	acgtagtaca	catggagata	ggagagcccg	atttagaacc	gtctcccaag	120
274	gtaatggaag	ctctggaacg	tgcggtgaag	gaaaagacgt	tcttctacac	ccctgctctg	180
276	ggactctggg	aactcaggga	aaggatatcg	gagttttaca	ggaaaaagta	cagcgttgaa	240
278	gtttctccag	agagagtcac	cgtaactacc	ggaacttcgg	gagcgtttct	cgtagcctac	300
280	gccgtaacac	taaatacggg	agagaagata	atcctcccag	acccctctta	cccctgttac	360
282	aaaaactttg	cctacctctt	agacgctcag	ccggttttcg	taaactgtga	caaggaaacg	420
284	aattacgaag	taaggaaaag	gatgatagaa	gacattgatg	cgaaaagccct	tcacatttcc	480
286	tcgcctcaaa	acctacggg	cacactctac	tcacctgaaa	ccctgaagga	acttgcgag	540
288	tactgcgaag	agaaggggtat	gtacttcata	tccgacgaga	tttaccacgg	actcgtttac	600
290	gaaggtaggg	agcacacagc	acttgagttc	tctgacaggg	ctattgtcat	aaacggggtt	660
292	tctaagtact	tctgtatgcc	aggtttcagg	atagggtgga	tgatagttcc	ggaagaactc	720
294	gtgagaaaag	cggaaatagt	aattcagaac	gtatttatat	ctgccccgac	gctcagtcag	780
296	tacgccgccc	ttgaggcttt	tgattacgag	tatttgagaa	aggtaaagaa	aacctttgaa	840
298	gagaggagga	acttccttta	tggggaaactg	aaaaaactct	tcaagataga	cgcgaaacct	900
300	cagggagctt	tttacgtatg	ggcaaacata	agtgattact	ccacagatag	ctacgaattt	960
302	gctttaaaac	ttttaaggga	ggcgaggggtg	gcggtaacgc	ccgggggtgga	ctttggaaaa	1020
304	aacaaaacga	aggagtatat	aaggtttgct	tatacgagaa	agatagaaga	acttaaggag	1080

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/389,537B

DATE: 04/04/2002  
TIME: 10:12:21

Input Set : A:\DIVER1240-3.ST25.txt  
Output Set: N:\CRF3\04032002\I389537B.raw

```

306 ggcgttgaaa ggataaagaa gttcttagag aagcttagct ga 1122
309 <210> SEQ ID NO: 19
310 <211> LENGTH: 1362
311 <212> TYPE: DNA
312 <213> ORGANISM: Aquifex
314 <400> SEQUENCE: 19
315 atgtgggaat tagaccctaa aacgctcgaa aagtgggaca aggagtactt ctggcatcca 60
317 ttaccacaga tgaaagtcta cagagaagaa gaaaacctga tatttgaacg cggagaaggc 120
319 gtttacctgt gggacatata cggcaggaag tatatagatg ccatactctc cctctggtgc 180
321 aacgtccacg gacataacca ccctaaactg aacaacgcag ttatgaaaca gctctgtaag 240
323 gtagctcaca caactactct gggaagttcc aacgttcccg ccatactcct tgcaaagaag 300
325 cttgtagaaa tttctcctga aggtattaaac aaggtctttt actccgaaga cggcgcgaa 360
327 gcagtagaga tagcgataaa gatggcttat cactactgga agaacaaggg agttaaggg 420
329 aaaaacgttt tcataacgct ttccgaagcc taccacgggg atactgtagg agcggtagc 480
331 gtagggggta tagaactctt ccacggaact tataaagatc tccttttcaa gactataaaa 540
333 ctcccatctc cttacctgta ctgcaaggaa aagtacgggg aactctgccc tgagtgcacg 600
335 gcagatttat taaaacaact ggaagatata ctgaagtcgc ggaagatat cgttgcggtc 660
337 attatggaag cgggaattca ggcagccgcg ggaatgctcc cttccctcc gggatttttg 720
339 aaaggcgtaa gggagcttac gaagaaatac gacactttaa tgatagttga cgaggttgcc 780
341 acgggatttg gcaggacggg aacgatgttt tactgtgagc aggaaggagt cagtcggagc 840
343 tttatgtgtc taggttaagg tataaccgga gggtagctcc cgcttgctgc gacactcaca 900
345 acggacgagg tgttcaatgc ctttttaggt gagttcgggg aggcaaagca cttttaccac 960
347 gggcacacct acactggaaa taacctcgcc tgttcggttg cactcgcaaa cttagaagtt 1020
349 tttgaggaag aaagaacttt agagaagtc caaccaaaga taaagctttt aaaggaaagg 1080
351 cttcaggagt tctgggaact caagcacgtt ggagatgta gacagctagg ttttatggct 1140
353 ggaatagagc tgggtgaagga caaagaaaag ggagaacctt tcccttacgg tgaaaggacg 1200
355 ggatttaagg tggcttacaa gtgcaggaa aaaggggtgt ttttgagacc gctcggagac 1260
357 gttatggtat tgatgatgcc tcttgtaata gaggaagacg aaatgaacta cgttattgat 1320
359 acacttaaat gggcaattaa agagcttgaa aaagaggtgt ag 1362
362 <210> SEQ ID NO: 20
363 <211> LENGTH: 1032
364 <212> TYPE: DNA
365 <213> ORGANISM: Aquifex
367 <400> SEQUENCE: 20
368 atgacatact taatgaacaa ttacgcaagg ttgcccgtaa agtttgtaag gggaaaagg 60
370 gtttacctgt acgatgagga aggaaaggag tatcttgact ttgtctccg tataggcgtc 120
372 aactccctcg gtcacgctta cccaaaactc acagaagctc taaaagaaca ggttgagaaa 180
374 ctccctccacg tttcaaactt ttacgaaaac ccgtggcagg aagaactggc tcacaaactt 240
376 gtaaaacact totggacaga agggaaggta tttttcgcaa acagcggaac ggaagtgta 300
378 gaggcggcta taaagctcgc aagggaagta tggagggata aaggaaagaa caagtggaa 360
380 tttatatcct ttgaaaactc tttccacggg agaacctacg gtagcctctc cgcaacggga 420
382 cagccaaagt tccacaaagg ctttgaacct ctagtctctg gattttctta cgcaaagctg 480
384 aacgatatag acagcgttta caaactccta gacgaggaaa ccgcggggat aattattgaa 540
386 gttatacaag gagagggcgg agtaaacgag gcgagtggag attttctaag taaactccag 600
388 gaaatttgta aagaaaaaga tgtgtcttta attatagacg aagtgcacac gggaaatagga 660
390 aggaccgggg aattctacgc atatcaacac ttcaatctaa aaccggacgt aattgcgctt 720
392 gcgaaggggac tcggaggagg tgtgccaata ggtgccatcc ttgcaaggga agaagtggcc 780
394 cagagcttta ctcccggtc ccacggctct accttcggag gaaacctctt agcctgcagg 840
396 gcgggaacag tggtagtaga tgaagttgaa aaactcctgc ctacgtaag ggaagtgggg 900

```

Use of n and/or Xaa has been detected in the Sequence Listing.  
Review the Sequence Listing to insure a corresponding  
explanation is presented in the <220> to <223> fields of  
each sequence using n or Xaa.

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/389,537B

DATE: 04/04/2002  
TIME: 10:12:22

Input Set : A:\DIVER1240-3.ST25.txt  
Output Set: N:\CRF3\04032002\I389537B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:35; N Pos. 986  
Seq#:36; Xaa Pos. 329